ĐẠI HỌC
BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY

f-Divergence Self-Play for Tabular Anomaly Detection via Large Language Models

Hoang Tran Vuong · Linh Ngo Van · Dang Nguyen · Thin Nguyen · Phuoc Nguyen · Mehrtash Harandi · Trung Le

Hanoi University of Science and Technology · Deakin University · Monash University

ICML 2026 — International Conference on Machine Learning

ONE LOVE. ONE FUTURE.

Content

1. **Introduction**
2. Background
3. Methodology
4. Experiments
5. Conclusion

1. Introduction

Anomaly Detection on Tabular Data

- **Goal:** Identify rare or abnormal instances that deviate from an underlying notion of *normality*
- **Challenge:**
  - Tabular heterogeneous data with non-trivial cross-field dependencies
  - Standard tabular AD pipelines assume fully structured features or rely on heavy feature engineering
- **Using LLMs:**
  - **How:** Operate on a serialized representation of each row; exploit semantics of categorical values and raw text with minimal preprocessing
  - **Challenge:** Tabular invariances, tokenizer-induced distortions, calibrated likelihood-based scores

AnoLLM (ICLR 2025)

- **Serialize** each row into a text template; discretize numerical values into bins; **fine-tune** on normal data via next-token prediction.
- **Inference:** score anomalies using negative log-likelihood.
- **Limitation:** One-shot SFT-style adaptation — weak signal for tightening the **boundary of normality**; performance plateaus on subtle, near-normal anomalies.
- **Motivation:** Reduce mismatch between the **model-induced distribution** and the **true normal-data distribution.**

1. Introduction

Self-Play and f-Divergence

● **Self-play:** Principled route to iterative refinement — contrast real data with the model's own generations via a discriminator-like objective; learning signal without additional annotations.

● **Tabular AD: Critic** distinguishes real normal samples from **model-generated** ones; **policy** updated so future generations better match the **normal-data distribution.**

● **Extension:** Broader family of ***f*-divergences** beyond SPIN's **IPM-like divergence** → useful when the geometry of the normal data manifold varies across domains.

Our Approach: DiSPaT

● **DiSPaT:** ***Divergence-driven Self-Play*** framework for unsupervised **Tabular** anomaly detection with LLMs.

● **Formulation:** Each mixed-type row is a serialized token sequence refined through an adversarial self-play game.

● **Objective:** Minimize the ***f*-divergence** between the **empirical distribution of normal data** and **the model-induced distribution.**

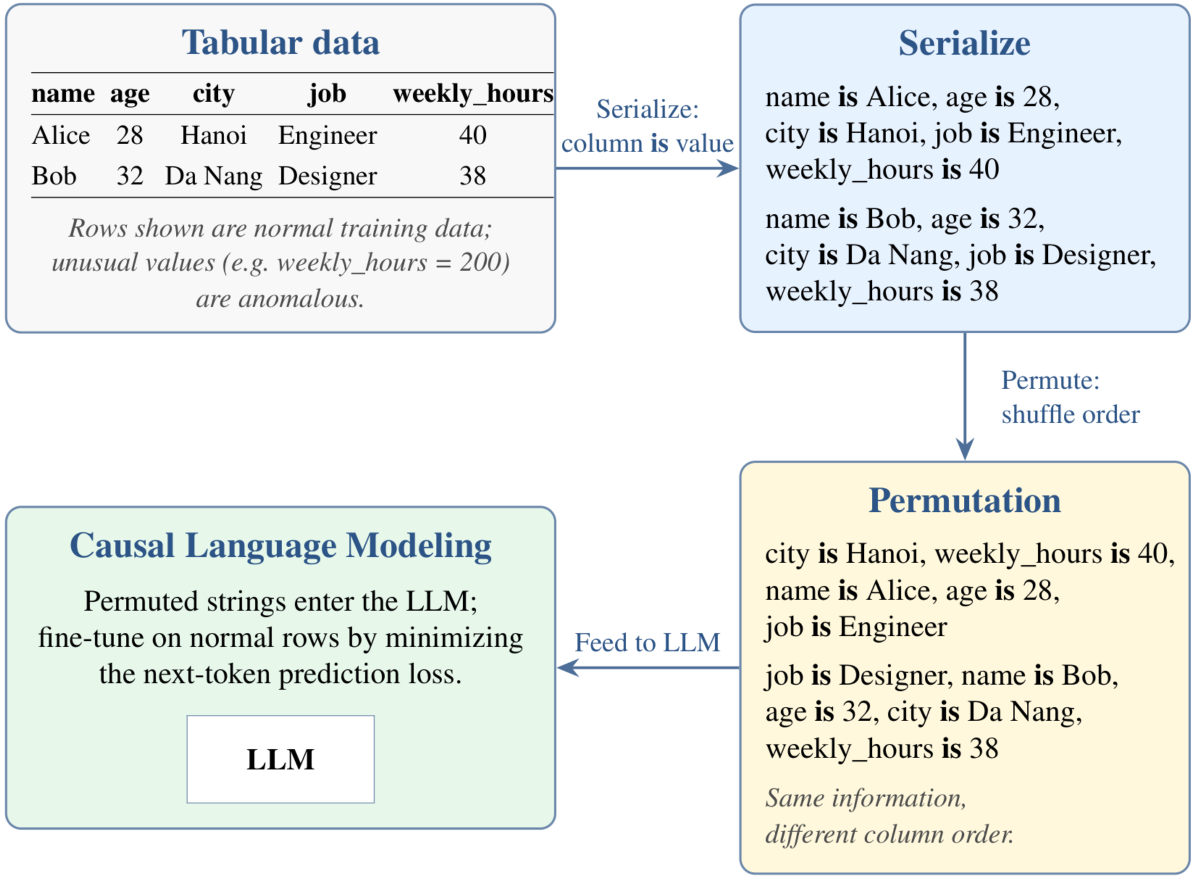
Content

2. Background: AnoLLM — Training

$$\mathbf{S}(\pi, \mathrm{x}_i) = \text{“}\mathrm{c}_{\pi(1)} \text{ is } \mathrm{E}(\mathrm{x}_{\mathrm{i},\pi(1)}), \ldots, \mathrm{c}_{\pi(\mathrm{d})} \text{ is } \mathrm{E}(\mathrm{x}_{\mathrm{i},\pi(\mathrm{d})})\text{”}$$

The resulting set of serialized tabular strings is denoted as $\mathcal{S} = \{\mathbf{S}(\pi, \mathbf{x}_i) \mid \pi \in S_d, i \in \{1, \ldots, n\}\}$. For each string $\mathbf{s} \in \mathcal{S}$, data is transformed into a token sequence $(\mathrm{w}_0, \mathrm{w}_1, \ldots, \mathrm{w}_{\mathrm{l}(\mathbf{s})}, \mathrm{w}_{\mathrm{l}(\mathbf{s})+1})$, where w_0 and $\mathrm{w}_{\mathrm{l}(\mathbf{s})+1}$ represent the beginning-of-sequence (BoS) and end-of-sequence (EoS) tokens, respectively. The fine-tuning process minimizes causal language modeling loss:

$$\mathcal{L}(\boldsymbol{\theta}) = \mathbb{E}_{\mathbf{s}\in\mathcal{S}}\left[-\sum_{k=1}^{l(\mathbf{s})+1} \log \pi_{\boldsymbol{\theta}}(\mathrm{w}_\mathrm{k} \mid \mathrm{w}_{0:\mathrm{k}-1})\right]$$

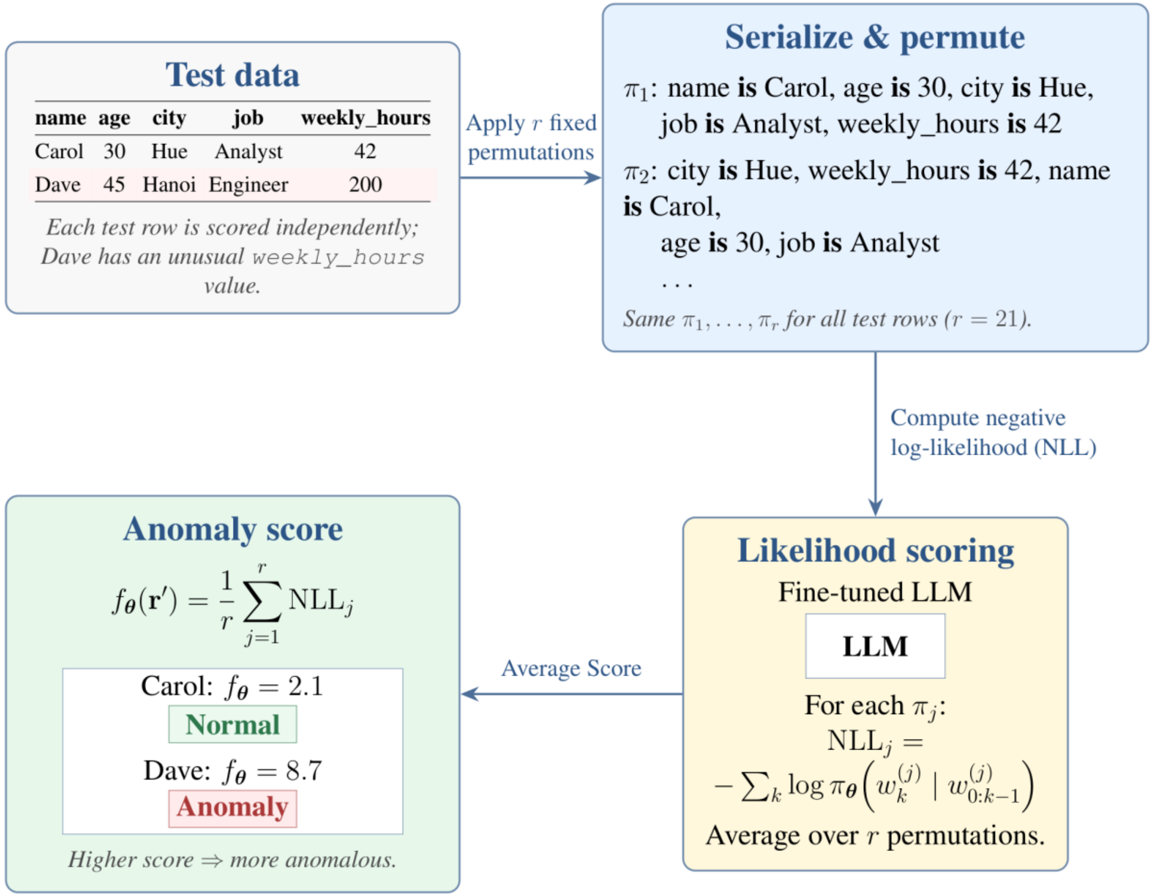


2. Background: AnoLLM — Inference

In the inference stage, anomaly score for a test sample $\mathbf{x}'$ is computed by averaging negative log-likelihood across r different permutations:

$$f_{\boldsymbol{\theta}}(\mathbf{x}') = -\frac{1}{r}\sum_{j=1}^{r}\sum_{k=1}^{l(\mathbf{S}(\pi_j,\mathbf{x}'))+1} \log \pi_{\boldsymbol{\theta}}(\mathrm{w}_k^{(j)} \mid \mathrm{w}_{0:k-1}^{(j)})$$

where $l(\mathbf{S}(\pi_j, \mathbf{x}'))$ denotes number of tokens used to encode tabular features under j-th permutation. Here, $\pi_1, \ldots, \pi_r \sim S_d$ denote distinct permutations of $\{1, 2, \ldots, d\}$, applied consistently across all samples.



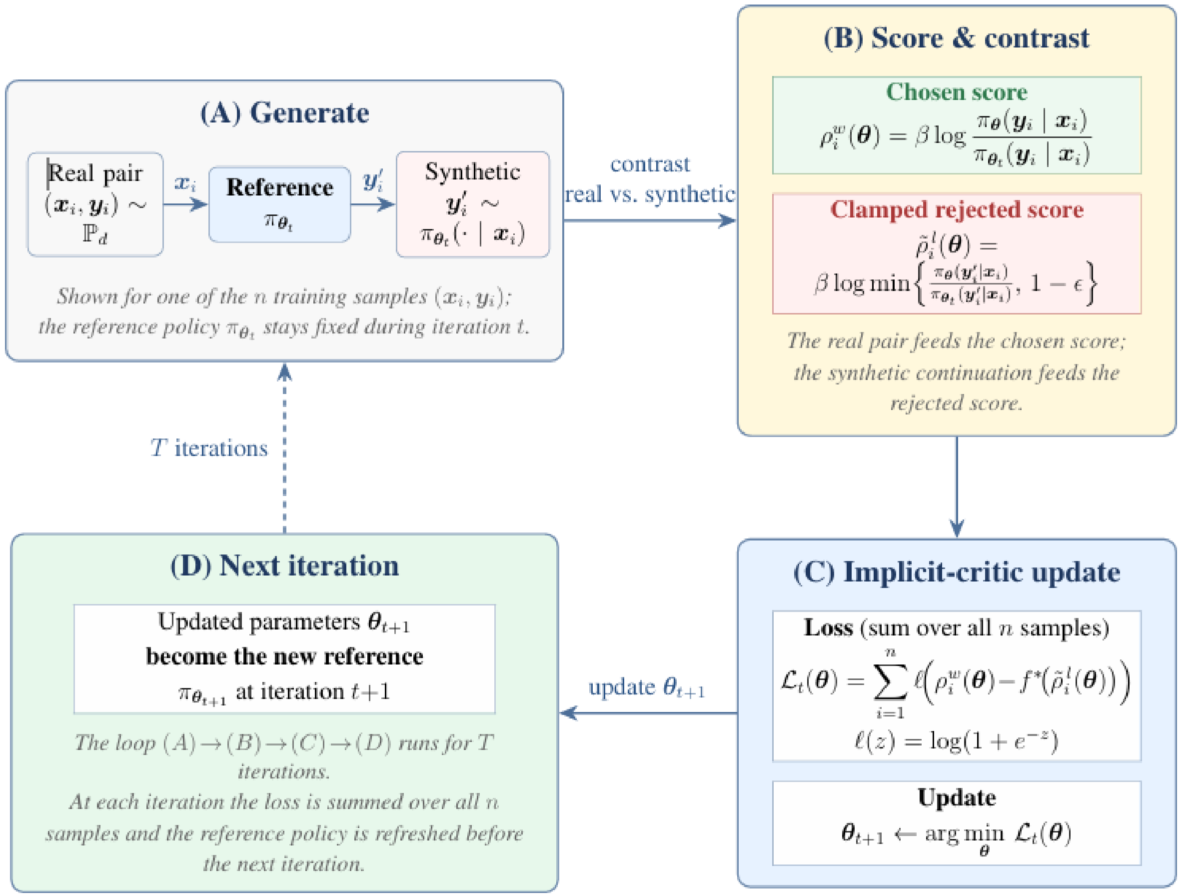
Content

3. Methodology

Let $\mathbb{P}_d$ denote the distribution over ground-truth normal sequences $(\boldsymbol{x}, \boldsymbol{y})$, where $\boldsymbol{x}$ is a context and $\boldsymbol{y} \sim \mathbb{P}_d(\cdot \mid \boldsymbol{x})$ is real continuation. Correspondingly, $\mathbb{P}_{\boldsymbol{\theta}}$ represents model-induced distribution with synthetic continuations $\boldsymbol{y}' \sim \pi_{\boldsymbol{\theta}}(\cdot \mid \boldsymbol{x})$. Our goal is to learn $\mathbb{P}_{\boldsymbol{\theta}}$ that closely aligns with $\mathbb{P}_d$ through iterative self-play

$$\min_{\boldsymbol{\theta}} D_f\left(\mathbb{P}_d, \mathbb{P}_{\boldsymbol{\theta}}\right)$$

$$D_f(\mathbb{P}_d, \mathbb{P}_{\boldsymbol{\theta}}) \approx \max_{T \in \mathcal{T}} \left\{ \mathbb{E}_{\mathbb{P}_d}\left[T(\boldsymbol{x}, \boldsymbol{y})\right] - \mathbb{E}_{\mathbb{P}_{\boldsymbol{\theta}}}\left[f^*\left(T(\boldsymbol{x}, \boldsymbol{y}')\right)\right] \right\}$$



3. Methodology

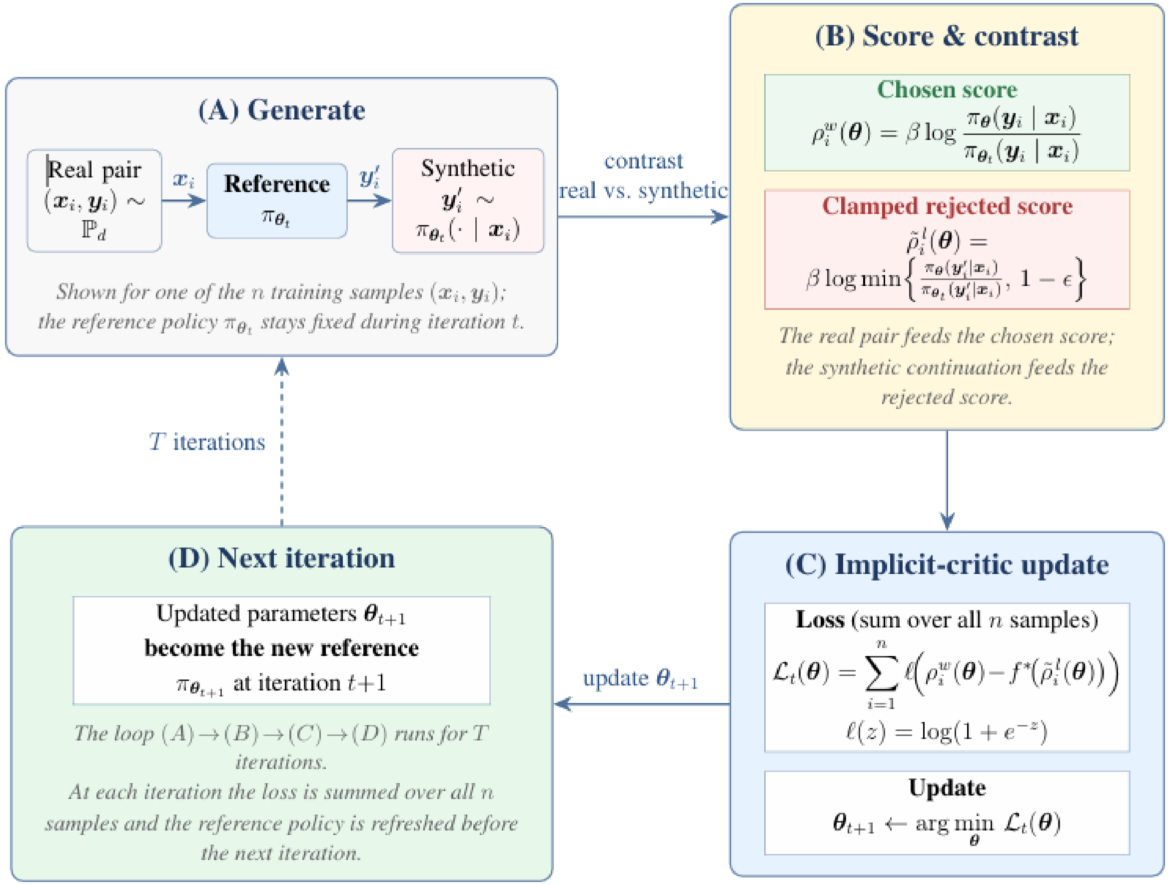
(1) Critic Update. Assume that at the current iteration k, we have current LLM model $\pi_{\boldsymbol{\theta}_k}$, and we update critic T_k to maximize:

$$\mathbb{E}_{\mathbb{P}_d}\left[T\left(\boldsymbol{x}, \boldsymbol{y}\right)\right] - \mathbb{E}_{\mathbb{P}_{\boldsymbol{\theta}_k}}\left[f^*\left(T\left(\boldsymbol{x}, \boldsymbol{y}'\right)\right)\right]$$

(2) Policy Update. Subsequently, we update LLM model $\pi_{\boldsymbol{\theta}}$ to move $\mathbb{P}_{\boldsymbol{\theta}}$ closer to $\mathbb{P}_d$ by performing the outer minimization, leading to:

$$\max_{\boldsymbol{\theta}} \mathbb{E}_{\mathbb{P}_{\boldsymbol{\theta}}}\left[f^*\left(T_k\left(\boldsymbol{x}, \boldsymbol{y}'\right)\right)\right]$$

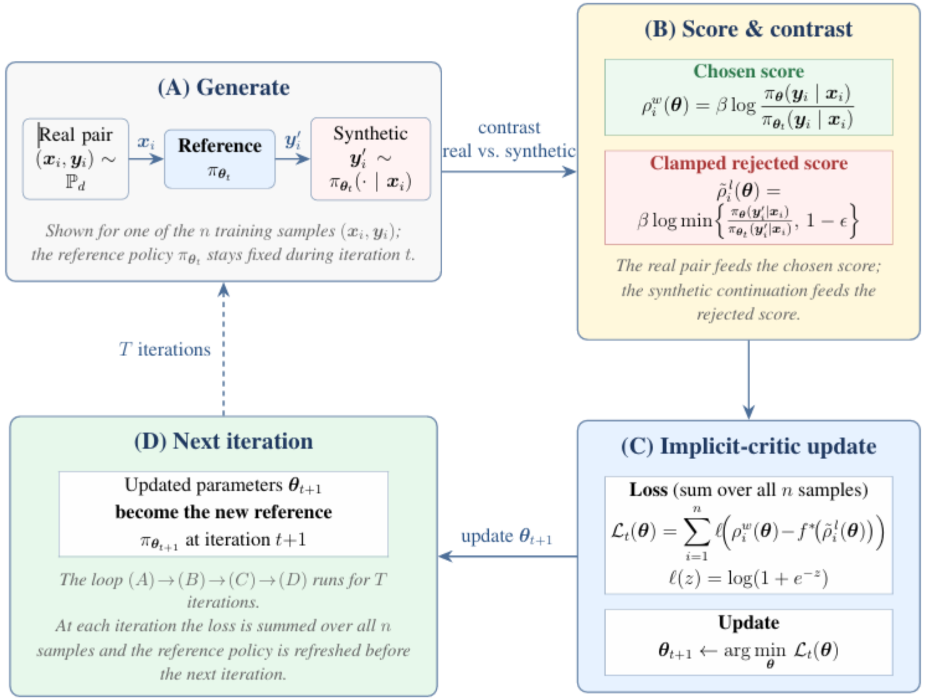
$$\min_{\boldsymbol{\theta}} \mathbb{E}_{\boldsymbol{x}}\left[\ell\left(\beta \log \frac{\pi_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x})}{\pi_{\boldsymbol{\theta}_k}(\boldsymbol{y} \mid \boldsymbol{x})} - f^*\left(\beta \log \frac{\pi_{\boldsymbol{\theta}}(\boldsymbol{y}' \mid \boldsymbol{x})}{\pi_{\boldsymbol{\theta}_k}(\boldsymbol{y}' \mid \boldsymbol{x})}\right)\right)\right]$$



3. Methodology

In practice, the model at iteration k may occasionally generate synthetic samples $\boldsymbol{y}'$ that closely resemble true normal data. Penalizing such samples as pseudo-anomalies would be counterproductive.To address this, we apply a clamping:

$$\tilde{\rho}^{l} = \beta \log \min \left\{ \frac{\pi_{\boldsymbol{\theta}}(\boldsymbol{y}'|\boldsymbol{x})}{\pi_{\boldsymbol{\theta}_k}(\boldsymbol{y}'|\boldsymbol{x})}, 1-\epsilon \right\}$$



Algorithm 1 DiSPaT

Input: Tabular dataset $\mathrm{X} = \{\mathbf{x}_i\}_{i=1}^{n}$, SFT-initialized LLM $\pi_{\boldsymbol{\theta}_0}$, iterations T, non-decreasing Fenchel conjugate f^*, regularization β, clamping threshold ϵ

for $i = 1, \ldots, n$ **do**
  Sample permutation $\pi \sim S_d$
  Serialize row $\mathbf{x}_i$ into context-continuation pair $(\mathbf{x}_i, \mathbf{y}_i) = \mathbf{S}(\pi, \mathbf{x}_i)$
end for
for $t = 0, \ldots, T-1$ **do**
  // Generate synthetic samples to serve as pseudo-abnormal
  for $i = 1, \ldots, n$ **do**
    Sample synthetic continuation $\mathbf{y}'_i \sim \pi_{\boldsymbol{\theta}_t}(\cdot \mid \mathbf{x}_i)$
  end for
  // Single implicit-critic update to minimize f-divergence
  Compute chosen reward: $\rho_i^w \leftarrow \beta \log \frac{\pi_{\boldsymbol{\theta}}(\mathbf{y}_i|\mathbf{x}_i)}{\pi_{\boldsymbol{\theta}_t}(\mathbf{y}_i|\mathbf{x}_i)}$
  Compute rejected reward with clamping:
$$\tilde{\rho}_i^l \leftarrow \beta \log \min \left\{ \frac{\pi_{\boldsymbol{\theta}}(\mathbf{y}'_i \mid \mathbf{x}_i)}{\pi_{\boldsymbol{\theta}_t}(\mathbf{y}'_i \mid \mathbf{x}_i)}, 1-\epsilon \right\}$$
  Update: $\boldsymbol{\theta}_{t+1} \leftarrow \mathrm{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^{n} \ell\left(\rho_i^w - f^*(\tilde{\rho}_i^l)\right)$
end for

Output: Trained model $\boldsymbol{\theta}_T$

Content

1. Introduction
2. Background
3. Methodology
4. **Experiments**
5. Conclusion

4. Experiments

Datasets

- 6 mixed-type: Fakejob, Lymphography, Seismic, Vehicle insurance, 20 Newsgroups, Ecoli
- 30 numerical (ODDS library) — Appendix

Backbone LLMs

- SmolLM-135M
- SmolLM-360M
- SmolLM-1.7B

Evaluation

- 50% normal for training; rest + anomalies for test
- Metrics: AUC-ROC, F1, AUC-PR Score
- Inference: r = 21 permutations

Baselines

Classical (4)

IForest · PCA · kNN · ECOD

Deep Learning (8)

DeepSVDD · RCA · SLAD · GOAD
NeuTral · ICL · DTE · REPEN

LLM-based (1)

AnoLLM

Recent SOTA — Appendix (5)

MCM · NPT-AD · DRL · LLM-DAS · CausalTAD

4. Experiments

Methods\Datasets	Fakejob	Lymphography	Seismic	Vehicle	20 Newsgroups	Ecoli	Average
			Classical methods				
Iforest	75.5%	67.3%	69.2%	49.6%	62.3%	85.6%	68.3%
PCA	72.4%	82.6%	69.2%	50.9%	62.3%	85.6%	70.5%
KNN	63.6%	86.0%	73.8%	52.4%	60.5%	87.7%	70.7%
ECOD	51.2%	83.0%	69.2%	50.9%	62.0%	77.6%	65.7%
			Deep Learning based methods				
DeepSVDD	56.1%	89.9%	71.3%	50.5%	59.7%	88.7%	69.4%
RCA	62.9%	91.9%	72.7%	53.1%	54.6%	88.3%	70.6%
SLAD	60.3%	96.4%	71.4%	55.6%	64.0%	88.2%	72.7%
GOAD	56.6%	81.7%	71.7%	51.2%	63.0%	88.1%	68.7%
NeuTral	54.8%	84.7%	68.1%	50.7%	65.8%	86.0%	68.4%
ICL	69.9%	82.7%	71.9%	50.1%	67.1%	88.7%	71.7%
DTE	54.8%	90.9%	71.4%	51.2%	60.0%	82.1%	68.4%
REPEN	65.3%	80.8%	72.4%	51.3%	57.4%	87.0%	69.0%
			LLM-based methods				
AnoLLM (SmolLM-135M)	80.0%	96.8%	71.2%	56.9%	76.6%	77.7%	76.5%
DiSPaT	84.2%$^{\uparrow}$	**100.0%**$^{\uparrow}$	73.2%$^{\uparrow}$	63.8%$^{\uparrow}$	**91.9%**$^{\uparrow}$	94.9%$^{\uparrow}$	84.6%$^{\uparrow}$
AnoLLM (SmolLM-360M)	81.4%	99.5%	74.6%	55.5%	75.2%	80.4%	77.8%
DiSPaT	**88.3%**$^{\uparrow}$	**100.0%**$^{\uparrow}$	**75.7%**$^{\uparrow}$	**64.2%**$^{\uparrow}$	87.6%$^{\uparrow}$	**96.8%**$^{\uparrow}$	**85.4%**$^{\uparrow}$
AnoLLM (SmolLM-1.7B)	80.2%	99.5%	74.0%	56.0%	77.4%	79.1%	77.7%
DiSPaT	85.2%$^{\uparrow}$	99.8%$^{\uparrow}$	74.6%$^{\uparrow}$	59.0%$^{\uparrow}$	87.1%$^{\uparrow}$	95.1%$^{\uparrow}$	83.5%$^{\uparrow}$

Table 1. Detailed AUC-ROC scores (%) comparing DiSPaT against baselines on the six-dataset benchmark.

4. Experiments

Methods\Datasets	Fakejob	Lymphography	Seismic	Vehicle	20 Newsgroups	Ecoli	Average
			Classical methods				
Iforest	27.4%	23.3%	25.1%	11.0%	13.7%	75.6%	29.4%
PCA	25.6%	56.7%	26.6%	12.4%	13.3%	**77.8%**	35.4%
KNN	16.3%	66.7%	29.1%	13.5%	15.6%	**77.8%**	36.5%
ECOD	16.5%	40.0%	28.2%	11.2%	13.2%	31.1%	23.4%
			Deep Learning based methods				
DeepSVDD	13.6%	56.7%	25.8%	11.5%	15.2%	53.3%	29.4%
RCA	13.7%	66.7%	32.0%	13.5%	12.9%	**77.8%**	36.1%
SLAD	17.5%	66.7%	28.5%	15.5%	15.9%	**77.8%**	37.0%
GOAD	12.9%	66.7%	29.5%	11.9%	13.6%	**77.8%**	35.4%
NeuTral	11.5%	63.3%	19.5%	12.0%	19.5%	51.1%	29.5%
ICL	24.5%	66.7%	29.8%	10.8%	19.0%	71.1%	37.0%
DTE	10.7%	66.7%	23.9%	12.1%	18.5%	66.0%	33.0%
REPEN	16.4%	66.7%	30.6%	12.6%	12.4%	75.6%	35.7%
			LLM-based methods				
AnoLLM (SmolLM-135M)	32.5%	76.7%	27.9%	16.2%	24.1%	33.3%	35.1%
DiSPaT	34.9%$^{\uparrow}$	**100.0%**$^{\uparrow}$	33.5%$^{\uparrow}$	**20.2%**$^{\uparrow}$	**49.8%**$^{\uparrow}$	55.6%$^{\uparrow}$	49.0%$^{\uparrow}$
AnoLLM (SmolLM-360M)	34.3%	80.0%	33.6%	17.4%	22.0%	33.3%	36.8%
DiSPaT	**40.8%**$^{\uparrow}$	**100.0%**$^{\uparrow}$	**35.9%**$^{\uparrow}$	19.6%$^{\downarrow}$	40.1%$^{\uparrow}$	**77.8%**$^{\uparrow}$	**52.4%**$^{\uparrow}$
AnoLLM (SmolLM-1.7B)	34.4%	83.3%	34.1%	16.5%	28.1%	44.4%	40.1%
DiSPaT	**37.2%**$^{\uparrow}$	83.3%	32.4%$^{\downarrow}$	18.4%$^{\uparrow}$	39.2%$^{\uparrow}$	**77.8%**$^{\uparrow}$	48.0%$^{\uparrow}$

Table 2. Detailed F1 scores (%) comparing DiSPaT against baselines on the six-dataset benchmark.

4. Experiments

Methods\Datasets	Fakejob	Lymphography	Seismic	Vehicle	20 Newsgroups	Ecoli	Average
Classical methods							
Iforest	22.7%	23.2%	23.5%	11.2%	14.6%	8.6%	17.3%
PCA	19.4%	62.4%	21.6%	11.7%	14.3%	16.0%	24.2%
KNN	13.8%	72.0%	25.6%	12.3%	14.8%	73.9%	35.4%
ECOD	13.0%	36.5%	24.4%	11.6%	14.1%	18.9%	19.8%
Deep Learning based methods							
DeepSVDD	12.0%	68.0%	22.6%	11.4%	13.5%	2.4%	21.7%
RCA	13.4%	78.3%	25.0%	12.4%	13.6%	17.6%	26.7%
SLAD	15.0%	79.5%	24.1%	14.0%	15.9%	11.0%	26.6%
GOAD	11.7%	69.7%	23.9%	11.6%	14.4%	1.2%	22.1%
NeuTral	10.8%	68.1%	19.3%	11.7%	17.6%	43.0%	28.4%
ICL	19.3%	71.8%	25.1%	11.5%	17.5%	66.4%	35.3%
DTE	10.6%	74.7%	22.4%	11.6%	15.7%	**77.7%**	35.5%
REPEN	14.6%	69.7%	24.9%	11.8%	12.6%	9.3%	23.8%
LLM-based methods							
AnoLLM (SmolLM-135M)	28.6%	85.6%	23.6%	14.1%	22.3%	20.6%	32.5%
DiSPaT	30.9%↑	**100.0%**↑	26.9%↑	**17.5%**↑	**49.4%**↑	53.6%↑	46.4%↑
AnoLLM (SmolLM-360M)	30.4%	93.8%	28.1%	14.3%	21.4%	12.7%	33.5%
DiSPaT	**36.3%**↑	**100.0%**↑	**29.1%**↑	16.9%↑	37.6%↑	65.4%↑	**47.6%**↑
AnoLLM (SmolLM-1.7B)	31.2%	97.6%	27.9%	14.3%	26.5%	39.4%	39.5%
DiSPaT	30.0%↓	97.6%	28.3%↑	15.1%↑	36.1%↑	77.3%↑	47.4%↑

Table 3. Detailed AUC-PR scores (%) comparing DiSPaT against baselines on the six-dataset benchmark.

4. Experiments

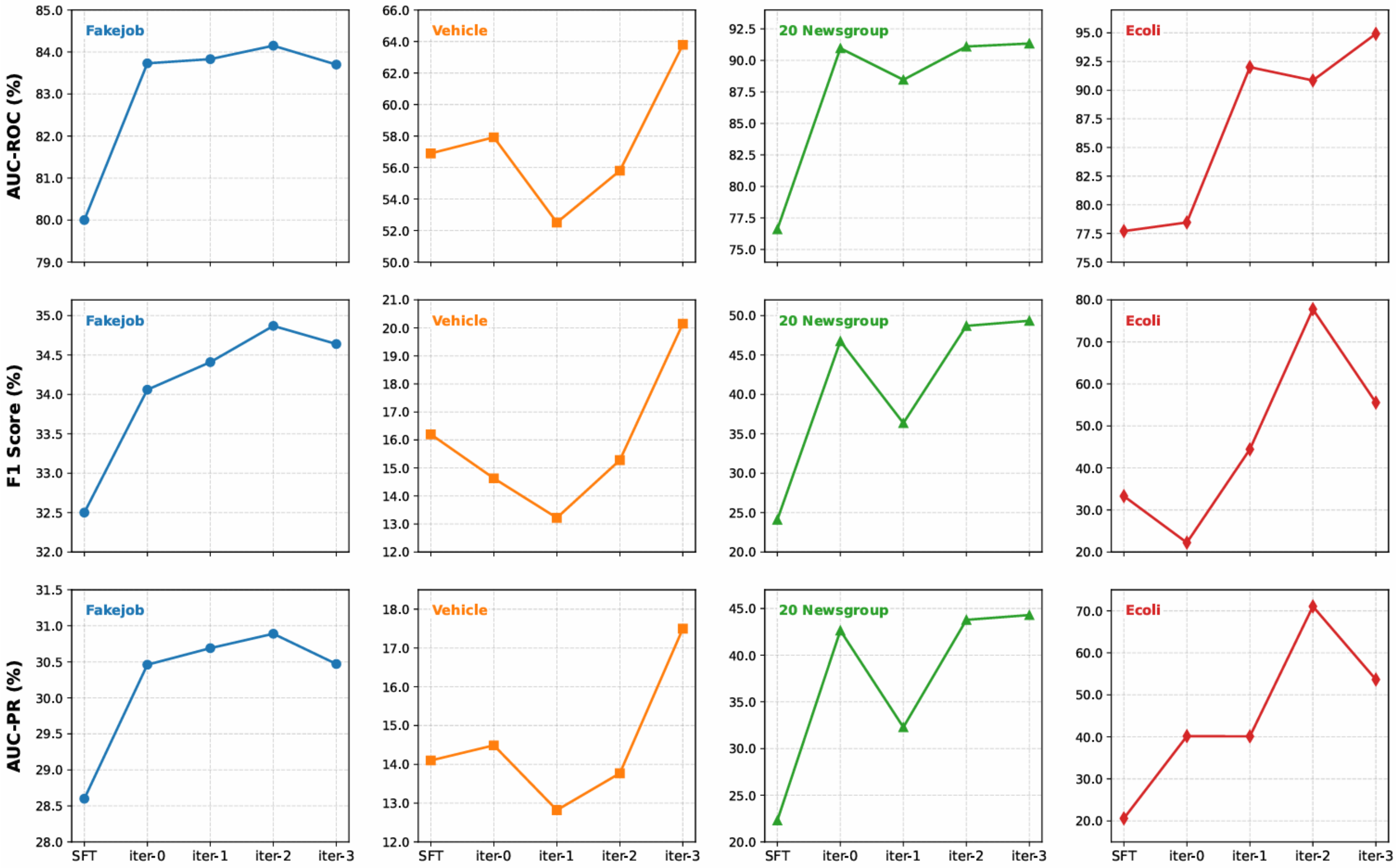


Figure 1. DiSPaT across self-play iterations on four datasets using SmolLM-135M. SFT indicates initial model, and iter-k refers to model after k iterations of training.

4. Experiments

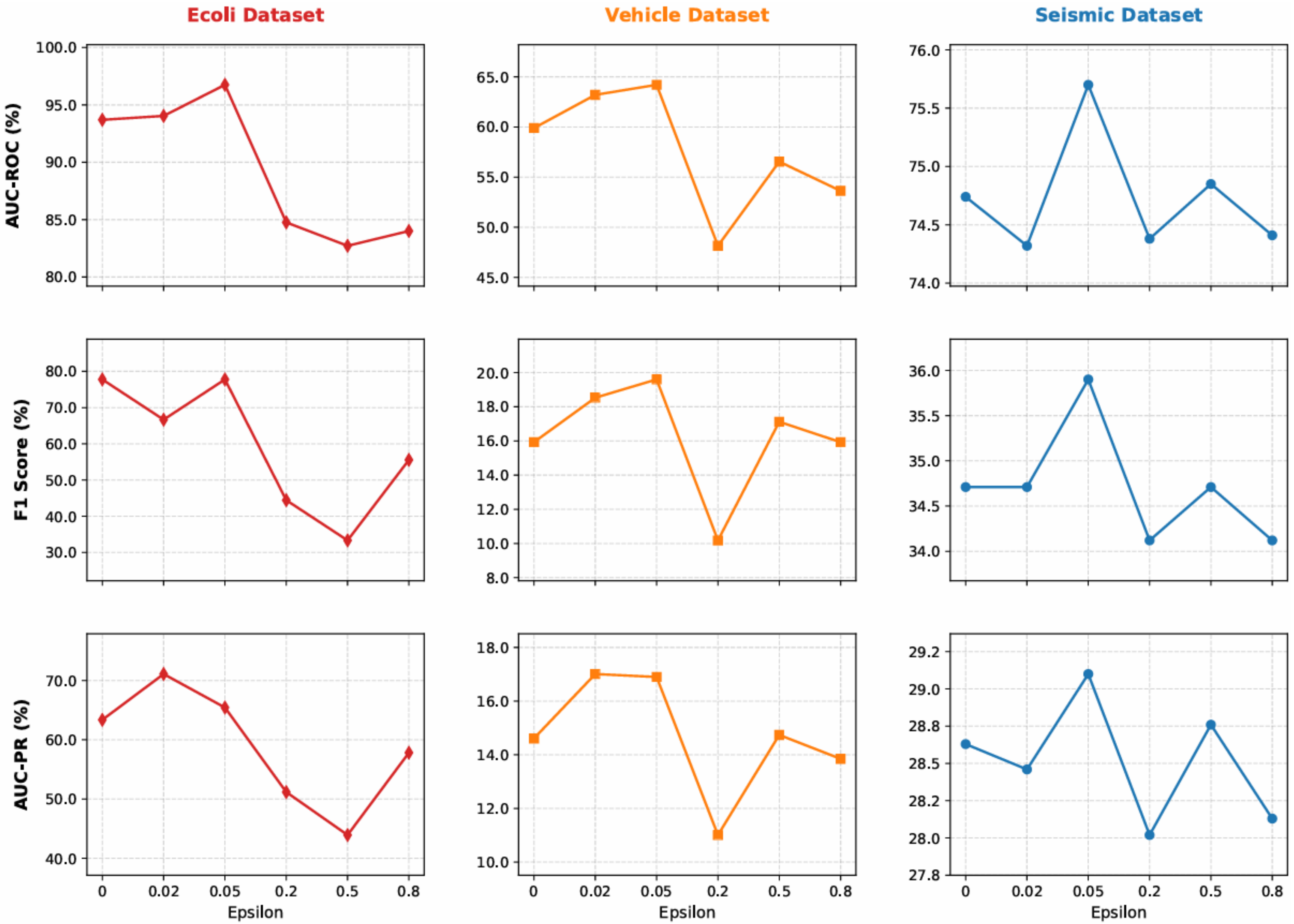


Figure 2. Effect of clamping threshold ϵ on three datasets usingSmolLM-360M.

4. Experiments

f^*	Metric	Fakejob	Lymphography	Vehicle insurance	Ecoli	Seismic	20 Newsgroups	Average
v	AUC-ROC	84.2%	99.3%	59.0%	92.3%	73.2%	91.3%	83,2%
	F1	34.9%	83.3%	16.9%	44.4%	32.4%	46.4%	43.1%
	AUC-PR	30.9%	94.4%	15.4%	37.9%	26.6%	46.7%	42.0%
$e^v - 1$	AUC-ROC	83.9%	100.0%	63.8%	94.9%	73.2%	91.9%	84.6%
	F1	35.1%	100.0%	20.2%	55.6%	33.5%	49.4%	49.0%
	AUC-PR	30.7%	100.0%	17.5%	53.6%	26.9%	49.8%	46.4%
$\frac{v}{1-v}$	AUC-ROC	84.0%	100.0%	60.5%	93.1%	73.0%	91.9%	83.8%
	F1	34.6%	100.0%	16.8%	55.6%	32.9%	48.6%	48.1%
	AUC-PR	30.9%	100.0%	15.6%	51.1%	27.0%	48.6%	45.5%
$-\log(1-v)$	AUC-ROC	83.8%	97.7%	62.8%	94.9%	73.1%	91.8%	84.0%
	F1	34.4%	83.3%	23.2%	55.6%	33.5%	48.6%	46.4%
	AUC-PR	30.6%	89.6%	23.2%	46.3%	26.4%	48.1%	44.0%

Table 4. Performance comparison of SmolLM-135M with different f-divergence choices across datasets.

4. Experiments

Dataset	AnoLLM Time	DiSPaT Time	AnoLLM FLOPs	DiSPaT FLOPs
Ecoli	4.1	28.5	1.5×10^{15}	3.7×10^{16}
Lymphography	4.4	30.9	3.1×10^{15}	1.7×10^{16}
20 Newsgroups	4.7	35.2	3.2×10^{15}	4.1×10^{16}
Vehicle	6.6	48.3	6.4×10^{15}	8.4×10^{16}
Seismic	9.7	67.6	9.8×10^{15}	1.2×10^{17}
Fakejob	98.8	671.3	8.4×10^{16}	7.7×10^{17}

Table 5. Training time and total FLOPs of AnoLLM and DiSPaT on six datasets using SmolLM-135M on a single NVIDIA A40 48GB GPU. Training time is reported in minutes.

4. Experiments

Benchmark	SPIN	DiSPaT
ARC	34.90 ± 1.39	$\mathbf{36.52 \pm 1.41}$
TruthfulQA	39.59 ± 1.45	$\mathbf{39.71 \pm 1.46}$
Winogrande	61.09 ± 1.37	$\mathbf{62.33 \pm 1.37}$
GSM8K	29.57 ± 1.26	$\mathbf{33.89 \pm 1.30}$
HellaSwag	43.66 ± 0.49	$\mathbf{45.11 \pm 0.50}$
MMLU	44.75 ± 10.05	$\mathbf{45.23 \pm 9.83}$
Avg.	42.26	**43.80**

Table 6. Comparison with SPIN in the standard alignment setting. We fine-tune Qwen1.5-1.8B on 50000 UltraChat200k prompts and evaluate on HuggingFace Open LLM Leaderboard benchmarks. Results are reported as mean and standard deviation.

4. Experiments

Datasets	Classical Methods				Deep-learning based Methods								LLM-based Methods					
	Iforest	PCA	KNN	ECOD	DeepSVDD	RCA	SLAD	GOAD	NeuTral	ICL	DTE	REPEN	135M		360M		1.7B	
													AnoLLM	DiSPaT	AnoLLM	DiSPaT	AnoLLM	DiSPaT
Annthyroid	92.2	83.9	81.1	79.0	74.2	71.8	76.1	57.2	81.3	84.2	97.7	73.6	92.7	92.9	93.1	93.1	93.0	94.0
Arrhythmia	82.7	79.6	78.6	81.1	76.5	78.6	78.4	68.1	76.0	78.5	77.1	68.4	82.5	81.4	82.2	84.3	82.4	82.4
BreastW	99.4	98.8	99.2	99.2	97.4	98.7	98.6	99.4	98.3	99.2	98.2	95.5	99.2	99.6	99.3	99.6	99.1	99.4
Cardio	94.8	96.6	92.1	93.5	84.2	94.8	84.0	52.4	85.9	89.4	92.0	82.9	94.0	94.9	87.3	91.7	86.7	87.7
Ecoli	85.6	85.6	87.7	77.6	88.7	88.3	88.2	88.1	86.0	88.7	82.1	87.0	77.7	94.9	80.4	96.8	79.1	95.1
ForestCover	87.0	94.5	98.5	92.1	53.3	94.4	85.7	27.8	89.8	97.7	97.8	90.2	88.1	92.2	83.5	91.3	88.7	88.9
Glass	80.2	71.3	84.9	69.3	82.4	71.9	79.0	57.4	93.3	88.7	79.9	75.5	81.9	82.6	79.7	82.3	81.8	82.5
Heart	81.9	84.2	81.4	65.9	77.3	78.5	82.2	83.8	81.1	78.7	83.1	44.9	82.0	79.7	79.9	80.7	80.3	83.3
Http	99.2	99.9	100.0	97.9	99.0	99.5	99.9	99.6	97.3	100.0	99.5	99.4	100.0	100.0	100.0	100.0	100.0	100.0
Ionosphere	89.1	89.4	96.0	73.4	96.3	91.6	96.0	95.0	95.6	97.0	96.4	54.5	90.9	96.4	92.4	92.7	91.8	93.0
LETTER	63.1	52.9	86.5	56.7	77.6	71.6	90.8	81.1	92.9	95.9	87.2	59.7	96.7	83.7	86.7	80.2	77.2	84.2
Lymphography	67.3	82.6	86.0	83.0	89.9	91.9	96.4	81.7	84.7	82.7	90.9	80.8	96.8	100.0	99.3	100.0	99.5	100.0
Mammography	88.1	90.0	87.2	90.6	85.7	87.3	74.0	75.6	69.0	78.2	86.4	86.3	91.5	87.2	87.6	88.4	87.4	88.1
Mulcross	99.9	100.0	100.0	96.0	100.0	100.0	96.9	100.0	96.8	100.0	100.0	97.3	100.0	100.0	100.0	100.0	100.0	100.0
Musk	97.1	100.0	100.0	95.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0	72.2	100.0	100.0	100.0	100.0	100.0	100.0
Optdigits	82.4	57.4	94.4	60.6	75.3	80.0	73.5	84.7	97.9	95.5	88.8	60.7	98.3	94.5	93.9	95.0	88.8	95.9
Pendigits	97.1	94.2	99.9	92.8	83.8	96.7	93.2	22.0	96.3	97.1	98.2	92.2	97.1	97.7	96.4	96.5	93.0	96.0
Pima	72.0	71.1	74.1	58.7	59.3	70.4	58.4	66.5	76.3	69.7	66.2	68.8	66.3	70.9	65.4	70.3	65.3	72.1
Satellite	80.7	66.2	87.4	58.2	80.0	73.4	86.4	79.5	85.9	85.4	78.9	74.0	90.2	88.7	87.7	90.2	85.8	89.8
Satimage-2	99.3	97.9	99.9	96.5	96.1	99.8	99.7	99.4	86.1	99.7	98.8	99.8	100.0	99.9	99.9	99.9	99.9	100.0
Seismic	69.2	69.2	73.8	69.2	71.3	72.7	71.4	71.7	68.1	71.9	71.4	72.4	71.2	73.2	74.6	75.7	74.0	74.6
Shuttle	99.6	99.4	99.9	99.3	99.7	99.6	99.8	99.0	99.7	99.9	99.8	99.2	100.0	99.9	100.0	99.9	99.9	100.0
Smtp	90.5	80.9	93.6	88.0	78.0	84.5	92.6	91.1	89.0	88.5	95.3	89.4	92.7	92.1	92.9	91.9	92.4	93.0
Speech	47.8	47.1	48.6	47.1	56.3	47.2	51.1	55.0	60.9	58.2	51.1	53.1	47.0	48.1	47.0	48.2	47.0	47.4
Thyroid	98.9	98.4	97.6	97.8	86.9	96.9	94.8	68.9	88.6	98.7	99.0	90.4	97.5	97.7	98.3	99.3	99.1	99.3
Vertebral	44.6	49.4	40.6	47.4	47.8	47.8	48.3	51.6	54.5	54.3	57.0	24.7	56.5	63.9	40.8	80.4	39.2	52.2
Vowels	77.9	64.4	97.5	59.7	89.5	89.1	96.9	92.5	98.8	97.9	98.3	75.3	98.2	96.0	93.8	95.5	93.3	94.9
WBC	94.7	94.9	94.7	90.7	90.4	94.6	92.8	66.3	76.5	90.8	91.3	77.0	96.4	94.8	95.2	95.5	95.7	95.0
Wine	93.0	92.7	95.2	72.9	85.4	89.9	96.4	96.1	96.0	94.4	97.4	83.5	90.9	97.2	85.1	96.2	87.6	96.0
Yeast	86.3	85.0	80.2	78.7	70.6	81.6	41.6	23.5	62.9	71.3	78.1	69.9	74.4	77.8	73.0	80.3	75.4	75.7
Average	84.7	82.6	87.9	79.0	81.8	84.8	84.1	74.5	85.5	87.7	87.9	76.6	88.4	89.4	86.5	90.5	86.1	89.4

Table 7. Detailed AUC-ROC scores comparing DiSPaT against baselines over 30 datasets in ODDS. The best and second-best results in each row are indicated in red and blue, respectively.

4. Experiments

Datasets	Classical Methods				Deep-learning based Methods								LLM-based Methods					
	Iforest	PCA	KNN	ECOD	DeepSVDD	RCA	SLAD	GOAD	NeuTral	ICL	DTE	REPEN	135M		360M		1.7B	
													AnoLLM	DiSPaT	AnoLLM	DiSPaT	AnoLLM	DiSPaT
Annthyroid	57.4	48.7	44.0	38.8	43.6	36.7	41.8	25.7	46.8	50.1	78.9	33.8	58.4	59.2	59.7	62.0	62.6	64.6
Arrhythmia	61.2	54.2	55.4	59.1	53.3	54.2	53.6	50.3	51.5	53.3	52.1	45.2	61.2	60.4	60.0	62.1	62.1	60.7
BreastW	96.9	95.9	96.3	95.4	93.6	95.9	95.1	96.6	96.7	95.9	96.3	93.5	95.8	97.1	96.6	97.1	96.7	96.7
Cardio	71.5	80.8	67.6	66.6	56.4	72.6	60.2	29.4	56.8	68.9	64.4	56.1	73.4	72.2	66.5	65.9	63.4	64.2
Ecoli	75.6	77.8	77.8	31.1	53.3	77.8	77.8	77.8	51.1	71.1	66.0	75.6	33.3	55.6	33.3	77.8	35.6	77.8
ForestCover	10.9	15.8	74.5	23.8	3.5	18.9	13.6	0.1	42.6	76.9	77.8	6.4	25.6	27.1	21.3	24.0	18.6	20.2
Glass	15.6	13.3	17.8	15.6	24.4	15.6	17.8	13.3	42.2	28.9	13.3	6.7	17.8	21.1	17.8	22.2	15.6	22.2
Heart	91.7	92.2	90.6	89.3	90.6	90.8	91.3	91.0	90.4	91.0	91.4	87.6	91.2	91.4	90.7	91.5	90.9	91.0
Http (KDDCUP99)	10.7	92.6	99.4	2.2	45.9	38.2	92.9	43.8	19.3	99.3	34.9	20.4	98.9	97.0	95.8	96.2	93.4	98.2
Ionosphere	79.7	78.9	89.4	66.0	89.5	83.2	89.4	85.4	88.2	90.6	90.0	59.5	82.1	89.7	83.8	84.4	84.3	85.7
Letter Recognition	17.6	13.6	43.4	14.6	40.4	29.0	54.8	40.4	63.6	72.2	58.8	16.4	73.4	66.2	48.6	47.4	32.0	40.1
Lymphography	23.3	56.7	66.7	40.0	56.7	66.7	66.7	66.7	63.3	66.7	66.7	66.7	76.7	100.0	80.0	100.0	83.3	83.3
Mammography	41.3	47.4	40.9	53.5	44.3	35.8	13.8	28.7	13.5	29.8	36.4	29.4	55.1	44.2	42.8	45.4	41.5	42.5
Mulcross	99.5	100.0	100.0	74.7	100.0	99.9	76.0	100.0	85.2	99.6	100.0	81.6	100.0	100.0	100.0	100.0	100.0	100.0
Musk	61.6	100.0	100.0	54.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0	15.0	100.0	100.0	100.0	100.0	100.0	100.0
Optdigits	15.9	0.1	28.4	2.7	29.7	2.1	4.0	16.5	63.9	47.1	16.4	2.0	72.0	50.7	44.3	46.7	34.0	63.3
Pendigits	55.1	44.2	91.0	42.7	44.5	53.0	35.6	0.0	46.7	61.0	60.6	37.3	55.9	57.1	50.5	52.6	36.7	56.4
Pima	67.2	68.8	69.2	57.8	57.0	67.2	58.5	62.8	69.5	67.0	62.4	66.8	62.6	64.6	62.0	66.0	61.7	68.3
Satellite	69.6	61.4	76.2	53.8	71.0	68.5	76.0	69.3	75.1	75.7	72.3	69.1	79.8	79.8	77.4	80.3	75.3	79.8
Satimage-2	87.3	84.8	93.5	71.8	88.4	94.9	82.5	95.5	5.1	91.8	50.1	91.6	95.2	94.4	94.4	94.4	92.6	95.8
Seismic	25.1	26.6	29.1	28.2	25.8	32.0	28.5	29.5	19.5	29.9	23.9	30.6	27.9	33.5	31.6	35.9	34.1	32.4
Shuttle	96.4	95.8	97.7	91.7	98.3	96.9	97.5	96.5	98.2	98.3	97.4	93.6	98.3	98.1	98.4	97.8	98.4	98.4
Smtp (KDDCUP99)	0.0	66.7	66.7	66.7	9.3	65.3	66.7	66.7	60.7	49.3	66.7	60.7	66.7	65.4	66.7	66.7	66.7	65.0
Speech	3.0	4.9	5.6	4.9	3.6	4.9	3.9	3.9	4.9	7.5	5.2	1.3	6.6	6.6	6.2	6.6	6.6	6.6
Thyroid	78.9	72.3	64.3	62.6	55.7	60.0	63.9	41.7	35.0	77.8	80.4	34.6	68.2	78.5	72.5	89.2	79.6	81.7
Vertebral	18.7	20.7	14.0	21.3	28.0	16.0	16.0	30.0	32.7	19.3	30.0	1.3	28.7	30.0	18.0	46.7	16.7	33.3
Vowels	24.8	20.0	68.4	22.0	54.8	42.4	71.6	46.0	76.8	74.0	78.0	24.4	76.0	64.8	55.6	60.0	60.0	66.0
WBC	81.9	79.0	71.4	56.2	70.5	72.4	66.7	38.1	18.1	73.3	61.9	21.9	79.0	78.4	76.2	85.7	71.4	76.2
Wine	70.0	62.0	70.0	38.0	42.0	60.0	72.0	76.0	70.0	62.0	72.0	44.0	50.0	80.0	52.0	70.0	62.0	70.0
Yeast	48.6	42.1	33.7	33.5	31.0	35.2	8.4	1.0	18.1	33.1	32.8	15.8	32.0	35.8	31.0	39.0	29.9	42.1
Average	51.9	57.2	64.8	46.0	53.5	56.2	56.6	50.8	53.5	65.1	61.2	43.0	64.7	66.6	61.1	67.8	60.2	66.1

Table 8. Detailed F1 scores comparing DiSPaT against baselines over 30 datasets in ODDS. The best and second-best results in each row are indicated in red and blue, respectively.

4. Experiments

Datasets	Classical Methods				Deep-learning based Methods								LLM-based Methods					
	Iforest	PCA	KNN	ECOD	DeepSVDD	RCA	SLAD	GOAD	NeuTral	ICL	DTE	REPEN	135M		360M		1.7B	
													AnoLLM	DiSPaT	AnoLLM	DiSPaT	AnoLLM	DiSPaT
Annthyroid	64.6	55.0	46.3	40.6	44.1	38.3	46.1	28.6	43.5	55.5	83.5	34.3	63.1	64.4	64.8	69.0	69.4	70.8
Arrhythmia	66.2	61.7	55.6	62.2	56.3	56.2	55.6	51.8	50.7	55.6	56.6	41.9	63.6	65.1	64.2	70.3	60.2	63.7
BreastW	99.4	98.5	99.2	99.2	96.6	98.6	98.3	99.4	97.0	99.1	96.7	92.5	99.1	99.6	99.2	99.6	99.2	99.5
Cardio	78.6	84.4	73.7	71.2	60.6	74.5	66.7	32.5	61.0	75.0	67.8	56.7	81.1	79.1	72.6	72.2	69.2	69.6
Ecoli	8.6	16.0	73.9	18.9	2.4	17.6	11.0	1.2	43.0	66.4	77.7	9.3	20.6	53.6	12.7	65.4	39.3	77.3
ForestCover	64.9	71.0	78.6	30.6	62.1	73.9	72.8	79.6	47.5	80.7	58.3	68.2	41.9	43.9	41.7	43.6	43.5	44.5
Glass	19.8	16.7	24.2	24.2	26.3	18.7	20.8	15.0	48.4	37.4	22.6	16.5	24.7	21.1	23.4	24.7	25.2	34.7
Heart	97.2	97.6	97.2	94.0	96.4	96.6	97.3	97.7	97.2	96.3	97.5	88.4	97.2	96.2	96.9	96.4	97.0	97.6
Http (KDDCUP99)	49.6	91.1	99.5	25.4	58.8	40.3	91.0	61.8	36.4	99.5	58.3	53.6	97.0	96.6	95.6	96.0	92.7	97.5
Ionosphere	89.8	91.2	96.7	76.9	96.7	93.2	96.9	95.8	95.9	97.7	97.2	53.4	93.3	97.3	93.2	94.5	94.2	94.9
Letter Recognition	16.8	14.3	42.6	14.1	41.2	25.9	57.8	39.0	70.3	77.3	55.0	15.1	79.7	50.2	19.1	46.9	29.2	45.7
Lymphography	23.2	62.4	72.0	36.5	68.0	78.3	79.5	69.7	68.1	71.8	74.7	69.7	85.6	100.0	93.8	100.0	97.6	97.6
Mammography	39.2	44.3	39.9	54.8	44.7	31.2	12.6	23.2	9.4	28.7	37.8	26.8	59.2	45.1	36.4	46.9	39.1	43.6
Mulcross	98.9	100.0	100.0	72.2	100.0	100.0	78.8	100.0	81.6	99.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Musk	66.6	100.0	100.0	62.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	17.5	100.0	100.0	100.0	100.0	100.0	100.0
Optdigits	16.6	5.9	31.4	6.5	23.2	12.2	10.9	17.8	64.5	41.4	22.2	7.6	75.0	58.1	39.8	43.5	28.8	66.4
Pendigits	54.4	37.6	95.8	39.5	41.6	51.6	29.2	2.6	40.8	65.6	50.9	31.9	62.3	67.7	55.4	56.9	42.5	63.4
Pima	71.4	69.6	71.6	62.2	60.6	69.8	60.3	66.0	74.6	69.5	63.9	67.3	67.7	70.1	67.4	68.2	66.3	70.4
Satellite	84.5	76.9	88.9	65.8	84.2	80.6	86.6	80.8	86.0	88.7	84.3	80.6	91.0	90.3	89.1	91.5	87.1	91.1
Satimage-2	93.0	90.1	98.0	74.5	90.5	97.7	90.3	98.0	8.2	96.7	52.6	95.2	98.8	98.1	97.4	98.1	96.5	98.8
Seismic	23.5	21.6	25.6	24.4	22.6	25.0	24.1	23.9	19.3	25.0	22.4	24.9	23.6	26.9	28.1	29.1	27.9	28.3
Shuttle	98.4	96.2	97.2	94.6	98.7	96.0	96.8	94.9	99.4	99.5	94.6	92.8	99.7	98.5	99.6	98.7	99.5	99.7
Smtp (KDDCUP99)	1.0	45.4	45.9	60.8	5.8	44.1	46.9	44.1	58.2	37.7	46.7	40.3	65.8	62.2	64.5	57.6	59.0	63.6
Speech	3.5	3.7	3.8	4.0	4.2	3.7	3.6	4.0	5.2	5.7	4.0	3.5	3.6	4.0	3.7	6.6	3.7	4.0
Thyroid	78.3	79.1	69.6	63.5	56.0	64.9	68.6	40.1	33.0	82.2	86.0	38.5	69.6	82.6	74.0	92.8	84.7	87.9
Vertebral	21.0	23.2	19.2	22.8	25.2	21.4	21.0	28.1	30.3	26.4	31.0	15.1	28.9	30.0	18.1	47.7	22.2	31.7
Vowels	22.9	16.2	76.2	15.3	60.3	45.5	76.5	54.4	86.1	80.4	83.1	20.3	83.9	64.7	59.9	60.0	61.4	68.0
WBC	84.2	87.6	81.4	58.6	74.7	80.8	71.1	40.8	22.6	71.4	64.0	23.5	87.3	83.5	75.3	89.0	75.8	81.4
Wine	67.2	65.9	71.1	32.1	51.2	51.7	78.2	78.9	77.9	73.4	87.3	48.4	52.2	87.3	52.9	78.1	68.1	80.3
Yeast	44.0	34.6	29.4	32.3	29.9	30.3	10.6	7.6	16.8	26.2	28.2	18.7	30.1	29.9	30.2	34.3	26.9	35.0
Average	54.9	58.6	66.8	48.0	56.1	57.3	58.7	52.6	55.8	67.7	63.5	44.4	68.2	68.8	62.3	69.3	63.5	70.2

Table 9. Detailed AUC-PR scores comparing DiSPaT against baselines over 30 datasets in ODDS. The best and second-best results in each row are indicated in red and blue, respectively.

Content

1. Introduction
2. Background
3. Methodology
4. Experiments
5. Conclusion

5. Conclusion

DiSPaT: Divergence-driven Self-Play for Tabular Anomaly Detection

- Unsupervised tabular anomaly detection with LLMs.
- Iterative f **-divergence minimization** – synthetic **pseudo-anomalies** at each iteration.
- Discriminative signal absent in one-shot fine-tuning → **tighter description of normality.**

Three Contributions

① Theory: **Minimise f-divergence** between normal data distribution and model's generations; tight variational approximation to the f-divergence between real and model-induced distributions.

② Algorithm: **DiSPaT** – self-play with implicit critic; single tractable optimization step per iteration; **no anomalous supervision.**

③ Empirical: **Consistently outperforms** AnoLLM, strong classical and deep-learning baselines, and recent tabular AD methods across multiple datasets; ablations validate iteration depth and divergence choice.